

ADATBÁNYÁSZAT A NEVELÉSTUDOMÁNY TERÜLETÉN (EDUCATIONAL DATA MINING)

R. Tóth Krisztina

SZTE BTK Neveléstudományi Doktori Iskola

Kulcsszavak: adatbányászat, pedagógiai értékelés, minőségfejlesztés

A technológia kínálja innovatív lehetőségek egyre nagyobb szerepet játszanak az oktatásban: előtérbe került a számítógéppel támogatott oktatás, tesztelő szoftverek, egyre elterjedtebb a web 2.0-ás eszközök tanórai használata. Ezeknek az eszközöknek, szoftvereknek az alkalmazásakor nagy mennyiségű adat áll elő automatikusan a tanulók tudásáról, attitűdjéről, tanulási stílusáról, és egyéb, a neveléstudományt érintő kérdésekről.

Az így keletkezett, háttérben automatikusan mentett adatok feldolgozása az adatok sokfélesége, a változók nagy száma miatt problémákba ütközhet mind a klasszikus, mind a valószínűségi tesztelméleti módszerek használatakor. A számítástudomány területéről származó adatbányászati algoritmusok lehetőséget kínálnak ezen problémák megoldására, az oktatáskutatás során alkalmazott módszerek finomítására és új modellek építésére. Az előadás célja az adatbányászat főbb alkalmazási területeinek, módszereinek bemutatása a neveléstudomány területén, gyakorlati példákkal illusztrálva.

Az adatbányászat (data mining, knowledge discovery in databases) nagyméretű adathalmazból összefüggések, mintázatok kinyerését jelenti (Romero és Yacef, 2007). Módszereit – az adott terület jellemzőihez adaptálva – számos kutatási területen alkalmazzák (pl.: bioinformatika, közgazdaságtan, nyelvészet, neveléstudomány). A neveléstudomány területén alkalmazott adatbányászat (Educational Data Mining, EDM) egy olyan terület, amely a tudományos vizsgálat középpontjába olyan módszerek fejlesztését helyezi, amelyek lehetővé teszik a tanulói teljesítmények mélyebb megértését a változók közötti komplex összefüggések elemzésével és a tanulási környezetekből kinyert egyedi adatokból újabb (pedagógiai szempontból releváns) tudományos következtetések levonását (Baker, 2009).

Az EDM fontosabb alkalmazási területei (Baker és Yacef, 2009): (1) a tanulói modellek tökéletesítése; (2) adott terület tudás-struktúra modelljének feltárása; (3) a tanítási/pedagógiai módszerek hatékonyságának vizsgálata; (4) empirikus bizonyíték keresése pedagógiai elméletek finomítására, kiterjesztésére, továbbá változók tanulásra gyakorolt hatásának mélyebb megértésére. Noha ugyanezen területeken alkalmazhatók az említett két tesztelméleti módszerei is, az EDM-algoritmusok alkalmazásával (pl. neurális hálók, döntési fák stb.) többszempontú, komplexebb elemzések végezhetők.

Az EDM használatára példa Beck és Mostow (2008) munkája, céljuk általános iskolai tanulók (n=346) szóolvasásának vizsgálata volt. A diákok a szóolvasást történetek (6,9 millió szó) olvasásával, oktatószoftver segítségével sajátították el. A rögzített adatok (pl. az olvasási idő, a félreolvasás és segítségkérés) felhasználásával exponenciális modellt építettek, majd a modell alapján, logisztikus regressziót alkalmazva, a tanulókat két csoportba tudták sorolni: az egyik csoport az adott történet újraolvasásával, a másik újabb történetek olvasásával tudta hatékonyabban elsajátítani a szóolvasást.

A kutatás a TÁMOP 3.1.9/08/01 kutatási program, az Oktatásméleti Kutatócsoport és az MTA-SZTE Képességkutató Csoport keretében zajlott.